Documentation du POC NPImageNearest

Cyril Vincent

**Introduction**

Utilisation des techniques d’IA pour rechercher les images d’un produit les plus proches des images d’autres produits dans NextPage.

Utilisation d’un modèle de Deep Learning CNN (Convolutional Neural Network) préentrainé de Google nommé FV (Feature Vector) et d’un modèle de Machine Learning nommé ImageHash inventé par Johannes Buchner et adapté par mes soins, le tout avec le framework Google TensorFlow 2.3.



**NPImageNearest**

Pour accéder à la démo, voici l’URL : <https://www.cyrilvincent.com/np/html/iindex.html>



Cette page liste l’intégralité des images du CHUV, avec l’id de l’image et la miniature de l’image.

Quand on clic sur l’image la page des plus proches images s’ouvre :



NPImageNearest trouve les 10 plus proches images avec un score > 82%

Signification des scores :

* >99% : c’est la même image
* >95% : c’est la même image légèrement retouchée
* >90% : c’est un produit très similaire
* >82% : image assez similaire
* <82% : non significatif

Essayons avec la seringue 266723



Nous remarquons que l’image 266708 est très proche mais il ne s’agit pas du même produit

J’ai essayé avec des images dupliquées et il les trouve avec un score de 100%

J’ai essayé avec des images recompressées et il les trouve avec un score de 99%

J’ai essayé avec des images retouchées et il les trouve avec un score de 90% à 95%

Dans le jeu de tests CHUV il n’y aucune image strictement identique, le tout sur 307 images

Signification de {'dah': 0.875, 'ddh': 0.781, 'dfv': 0.942, 'dsize': 47, 'dn': 0.652} :

* Dah est le score sur 1 du modèle AverageHash de ImageHash, il s’agit d’une hyper compression de l’image en 8x8 puis d’une mesure de distance entre les 2 images compressée. Ce modèle est très rapide mais donne des faux positifs
* Ddh est le score sur 1 du modèle DifferenceHash de ImageHash qui fonction comme AverageHash mais sur les gradients de luminosités. Ce modèle est rapide et détecte très bien les images retouchées
* Dfv est le score sur 1 du modèle Feature Vector qui est un réseau Deep Learning CNN préentrainée sur ImageNet qui va hacher l’image sur 1792 bits. Ce modèle va trouver une même forme sur des images pourtant différentes. Ce modèle est assez lent en apprentissage mais rapide en prédiction
* Dsize est la différence entre la signature de l’image, ce modèle détecte pour un coût dérisoire 2 images strictement identiques
* Dn n’est pas utilisé
* Le score est calculé par une moyenne pondérée des différents scores après un léger apprentissage.

**Conclusion**

L’outil est fiable à 100% pour retrouver des images strictement identiques

L’outil est fiable à 95% pour retrouvées des images recompressées ou retoucher

L’outil est fiable à 95% pour retrouver des produits identiques sur un ensemble de produits assez discriminants les uns aux autres

L’outil est fiable à 90% pour retrouver des produits identiques sur un ensemble de produits où les produits se ressemblent tous comme pour CHUV

**Efficacité**

Une prédiction par apprentissage profond nécessite environ 32 secondes pour 1000 images et 50 minutes pour 100000 !

J’ai donc utilisé une technique d’indexation NoSql pour améliorer les performances.

J’ai généré des fichiers fictifs avec 10000 et 100000 images pour tester les temps de traitement.

Voici les performances en prédictions des produits proches sur mon PC perso (i7 récent + GPU + 16Go de RAM) :

* Pour 305 images : <0.1s
* Pour 1000 images : 0.1s
* Pour 100000 images : 5s

Le temps d’apprentissage est le suivant :

* Pour 305 images : 0.1s
* Pour 1000 images : 0.2s
* Pour 100000 images : 17.5s

Voici les performances de l’apprentissage pour comparer N images parmi M images en désactivant le modèle FV pour accélérer l’apprentissage et en baissant un peu la qualité de prédiction :

* Pour 305x305 images : 7s
* Pour 1000x1000 images : 74s
* Pour 5000x5000 images : 51 minutes

**Détail d’utilisation du modèle CNN FV et ImageHash**

* Clonage du modèle CNN FV entrainé depuis MobileNet
* Adaptation des poids par transfert learning
* Création de l’indexation des caractéristiques en JSON et Pickle
* Hachage des caractéristiques par prédiction du modèle avec création d’un vecteur de 172 flottants 64 bits et 2x64 flottants 64 bits pour AH et DH
* Stockage des hachages en index
* Recherche des plus proches voisins par application un calcul de distance avec similarité cosine
* Modification et normalisation des poids

**Industrialisation et intégration dans NextPage**

Ceci est un POC, mais il est très facile d’industrialiser le code

Pour ceci il suffit d’encapsuler l’apprentissage et la prévision soit dans un service REST multi-process.

NP pourra facilement l’interroger par requêtes REST + JSON, soit par un wrapper .NET, un petit transfert de compétence doit avoir lieu pour installer et utiliser le service

Coût : 2j

**Opportunités**

Il est possible de poursuivre la démarche du POC pour améliorer NextPage :

* Catégorisation : Trouver automatiquement la famille d’une image
* Prédiction : prédire un produit lié ou une caractéristique
* Reporting : Afficher tous les produits similaires entre eux à plus de 90%
* Barcode : Lire automatiquement le code barre sur l’image

Ces fonctionnalités peuvent se faire en POC entre 2 jours chacune.

**Références**

TensorFlow : <https://www.tensorflow.org/?hl=fr>

Feature Vector : <https://tfhub.dev/google/tf2-preview/mobilenet_v2/feature_vector/4>

ImageHash : <https://pypi.org/project/ImageHash/>